Poster: Feasibility of Malware Traffic Analysis through TLS-Encrypted Flow Visualization

IEEE International Conference on Network Protocols 2020
October 13-16, 2020

Dongeon Kim, **Jihun Han**, Jinwoo Lee, Heejun Roh

Korea University Sejong Campus, Sejong, Republic of Korea

Wonjun Lee

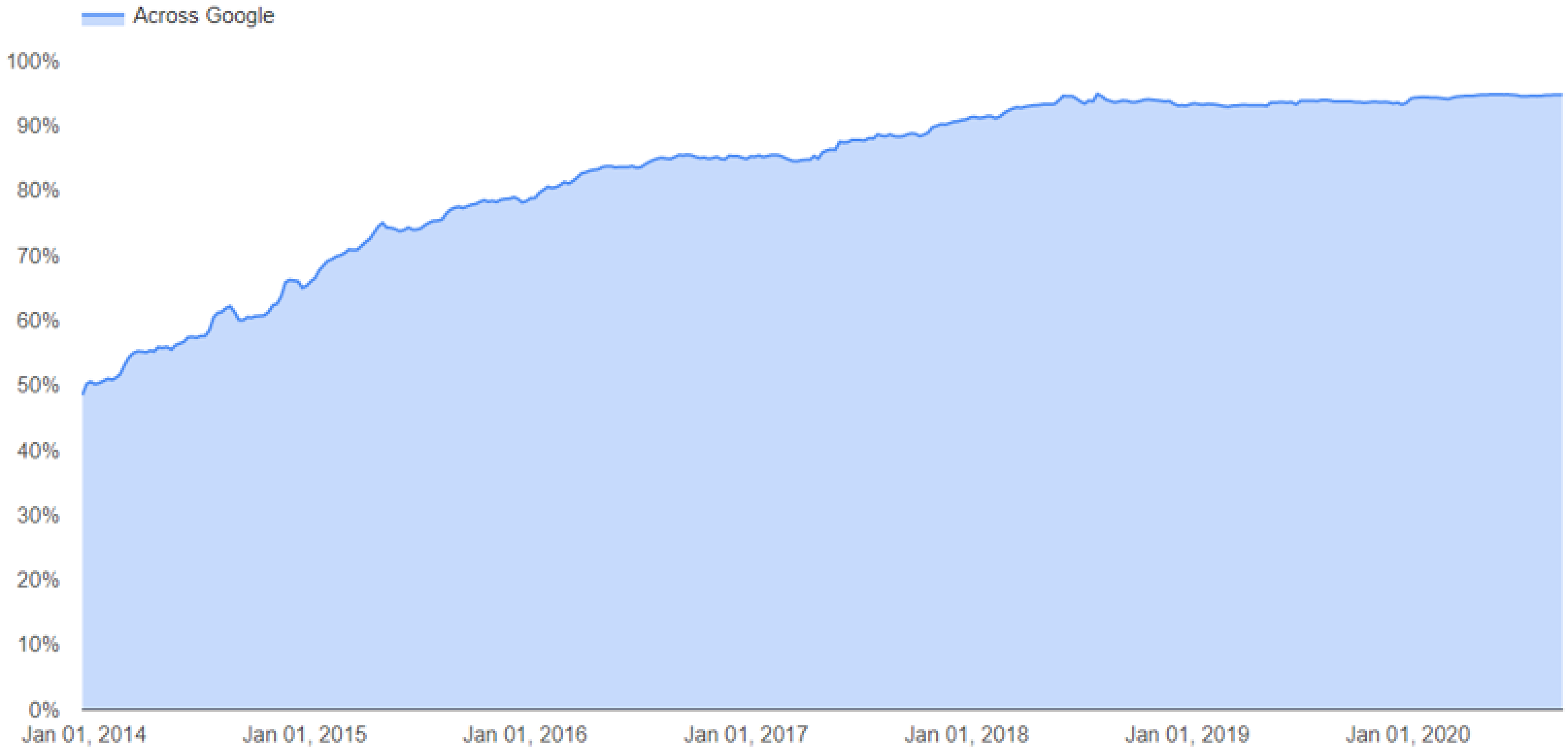
Korea University, Seoul, Republic of Korea

KOREA UNIVERS
SEJONG CAMPUS

KOREA
UNIVERSITY

Motivation

Encrypted traffic across google



Network using TLS encryption is increasing

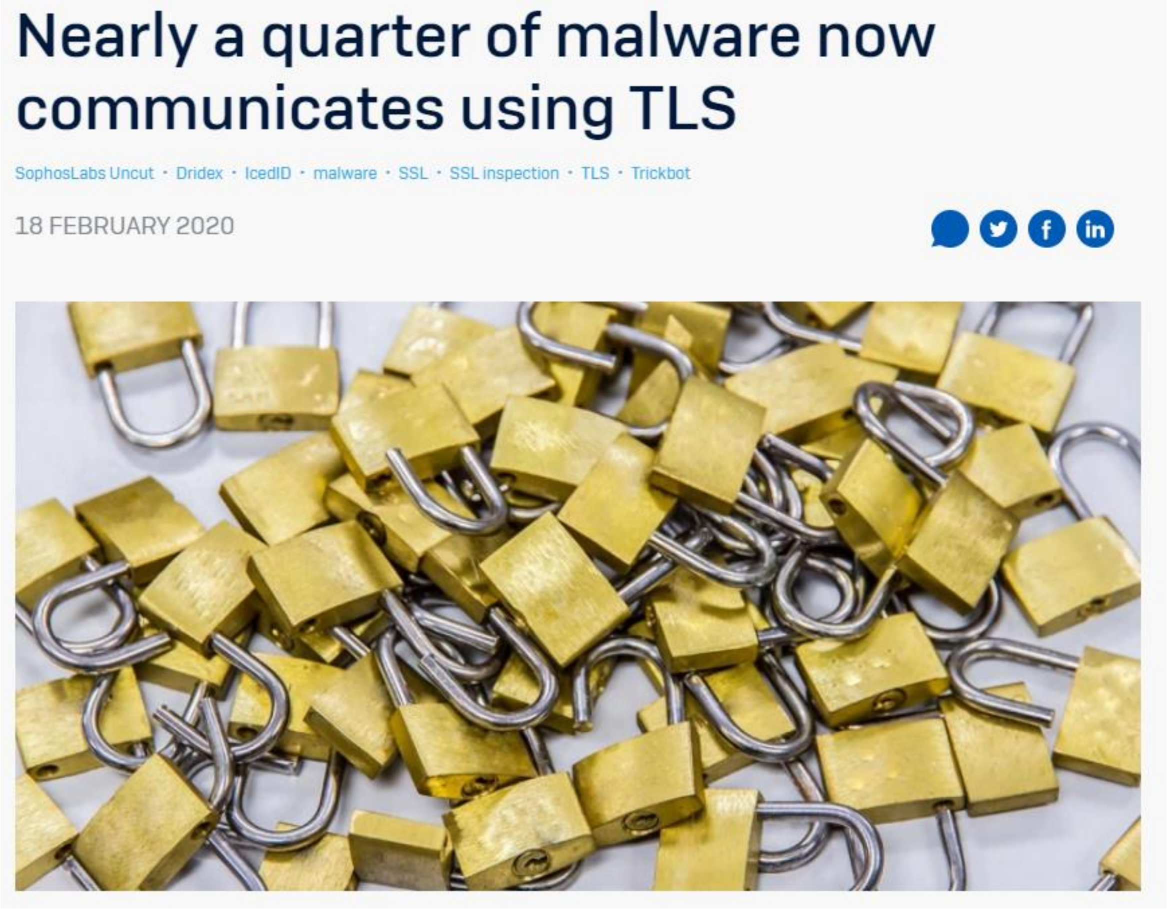
95% of traffic across google is encrypted

80% of enterprise traffic on the Zscaler cloud in is encrypted

KOREA UNIVERSITY SEJONG CAMPUS

https://transparencyreport.google.com/https/overview?hl=en

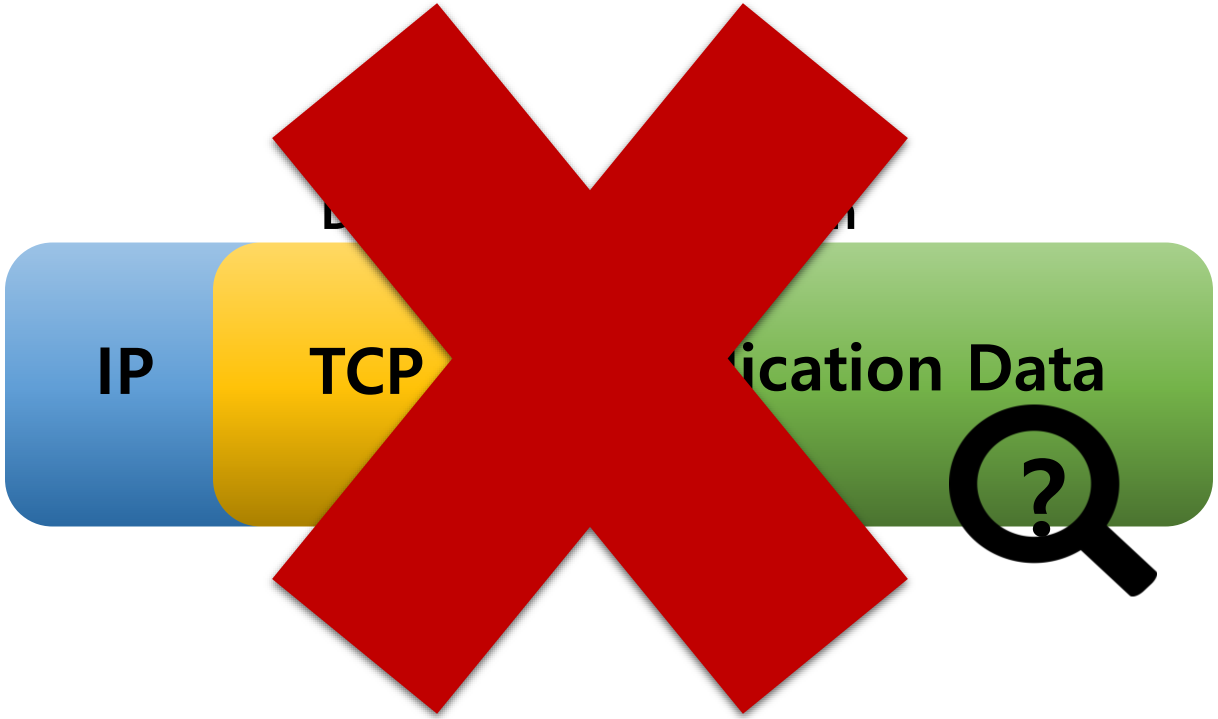
Motivation



Nearly a quarter of malware now communicates using TLS

SophosLabs Uncut · Dridex · IcedID · malware · SSL · SSL inspection · TLS · Trickbot

18 FEBRUARY 2020



KOREA UNIVERSITY SEJONG CAMPUS

https://news.sophos.com/en-us/2020/02/18/nearly-a-quarter-of-malware-now-communicates-using-tls

Motivation

- B. Anderson and D. McGrew, “**Identifying encrypted malware traffic with contextual flow data,**” in *Proc. of AISec’16 (co-located with ACM CCS)*, Vienna, Austria, October 2016.
- B. Anderson, S. Paul, and D. McGrew, “**Deciphering malware’s use of TLS (without decryption),**” *Journal of Computer Virology and Hacking Techniques*, vol. 14, no. 3, pp. 195–211, August 2018.

- **Require fine-grained feature selection conducted by experts**
- **Need to conduct field-specific preprocessing for message field values**

Our Proposal: TLS-Encrypted Flow Visualization

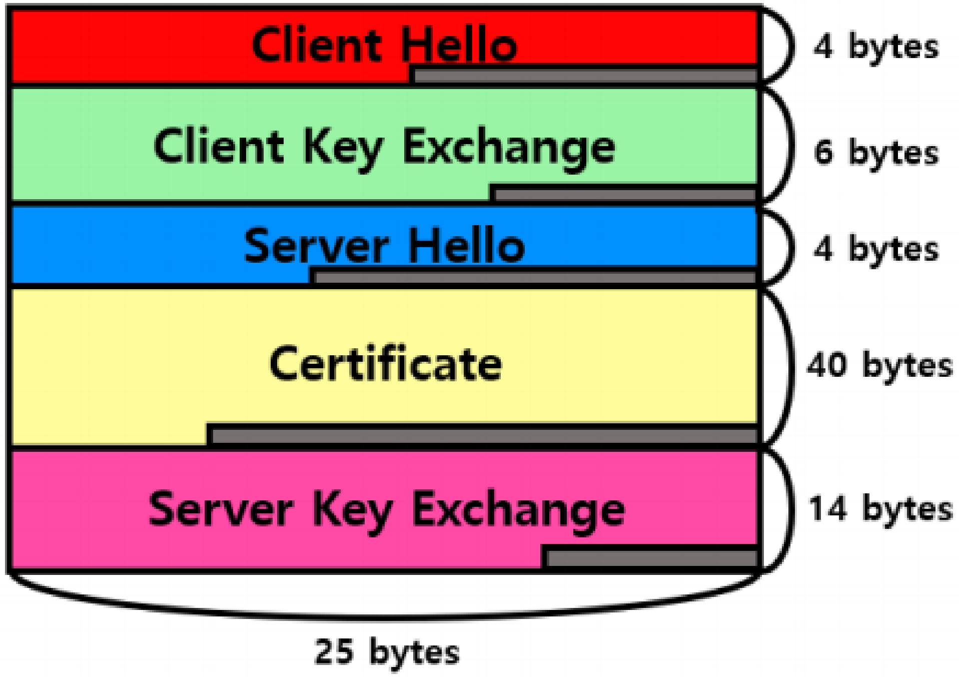


Image Format of TLS-Encrypted Flow

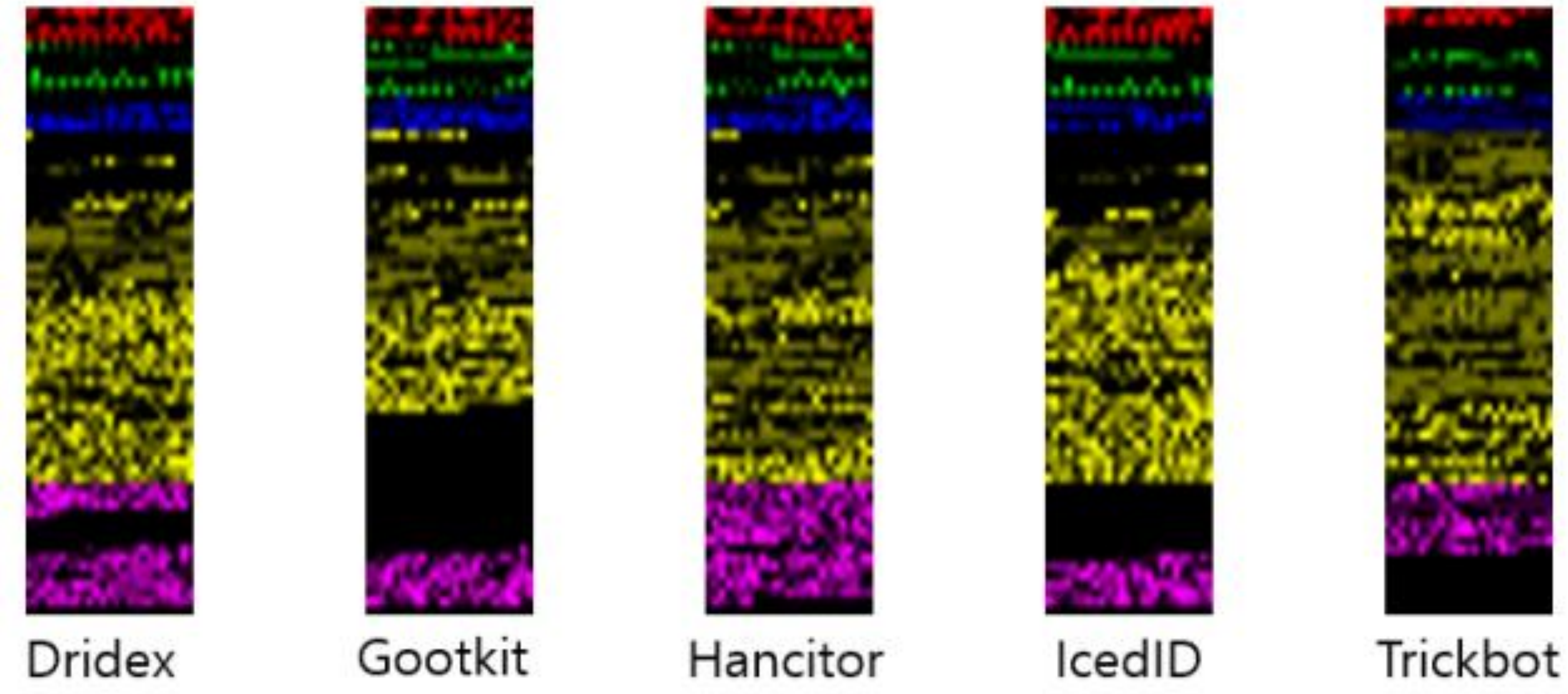
Our Proposal: TLS-Encrypted Flow Visualization

- TLS flow metadata have fruitful information to classify encrypted malware traffic

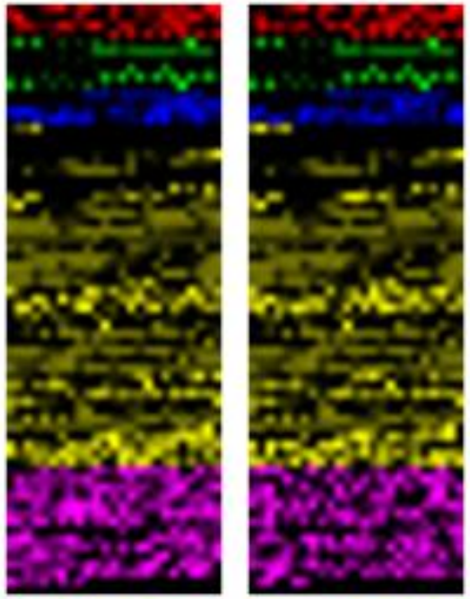
- Images can capture small changes yet retain the global message exchange pattern

- Different messages of a flow can be easily observed as a colored image

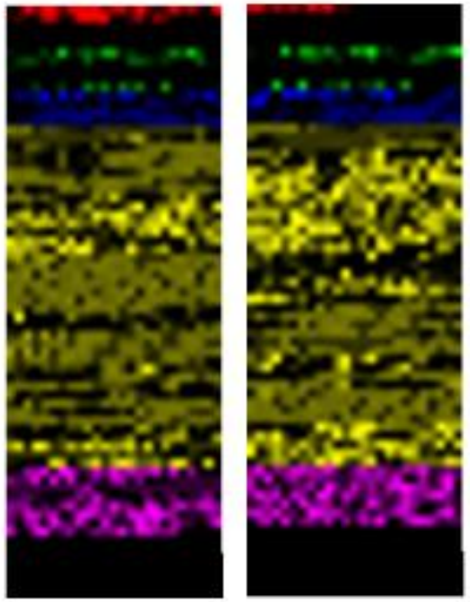
Images from Malware Families



Feasibility of Malware Traffic Analysis via Images

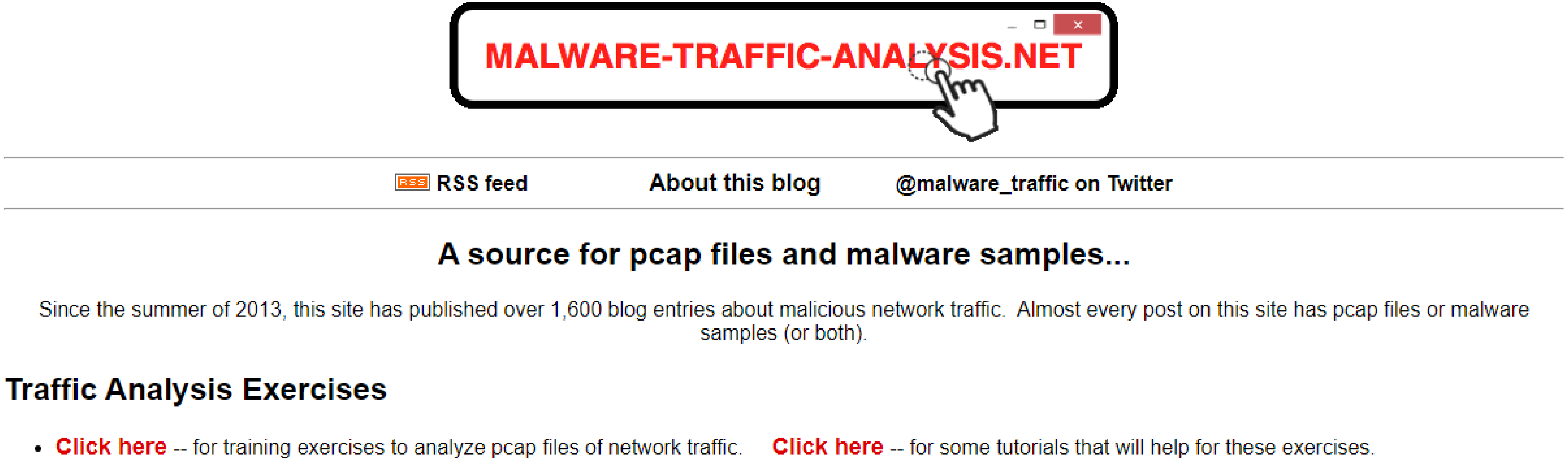


(a) Hancitor Sample Images



(b) Trickbot Sample Images

Experimental Results



MALWARE-TRAFFIC-ANALYSIS.NET

RSS feed About this blog @malware_traffic on Twitter

A source for pcap files and malware samples...

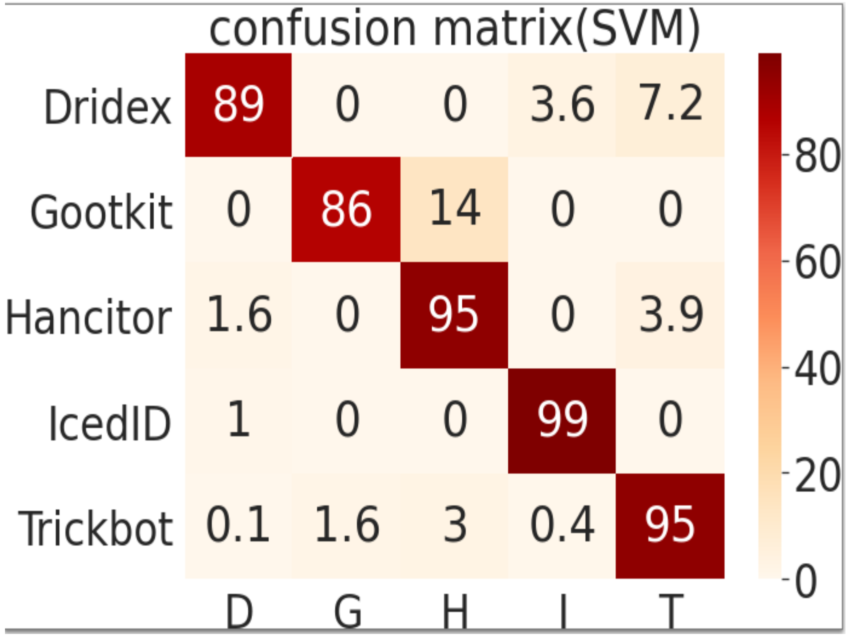
Since the summer of 2013, this site has published over 1,600 blog entries about malicious network traffic. Almost every post on this site has pcap files or malware samples (or both).

Traffic Analysis Exercises

- Click here -- for training exercises to analyze pcap files of network traffic. Click here -- for some tutorials that will help for these exercises.

B. Duncan. Malware traffic analysis. [Online]. Available: http:/malware-traffic-analysis.net/

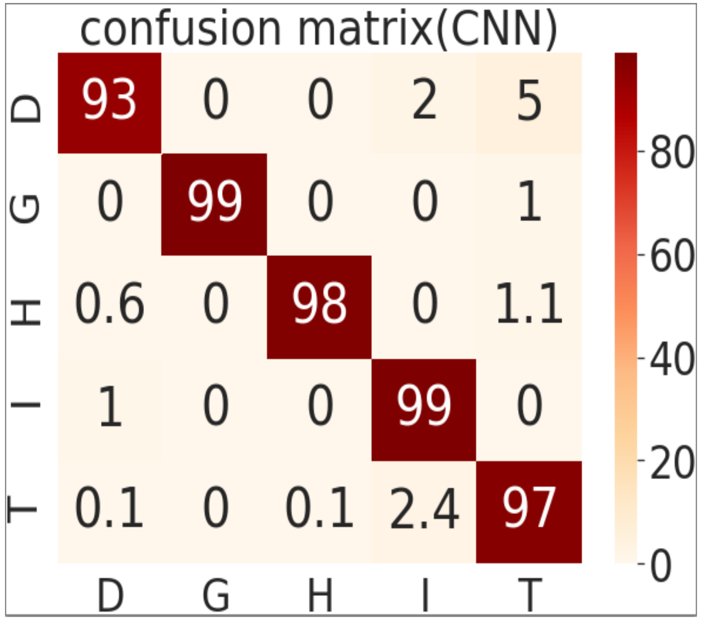
Experimental Results



confusion matrix(SVM)

	D	G	H	I	T
Dridex	89	0	0	3.6	7.2
Gootkit	0	86	14	0	0
Hancitor	1.6	0	95	0	3.9
IcedID	1	0	0	99	0
Trickbot	0.1	1.6	3	0.4	95

93% Accuracy in Average



confusion matrix(CNN)

	D	G	H	I	T
D	93	0	0	2	5
G	0	99	0	0	1
H	0.6	0	98	0	1.1
I	1	0	0	99	0
T	0.1	0	0.1	2.4	97

97% Accuracy in Average

Resulting confusion matrices

Conclusion

- Malware using TLS will continue to increase in the future
- There needs to be new method to detect malware using TLS
- Both SVM and CNN had high accuracy, even though the images do not have similar patterns